

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-299829

(43)Date of publication of application : 24.10.2000

(51)Int.Cl.

H04N 5/76
G06F 17/30
G06T 1/00
H04N 5/91
H04N 5/765

(21)Application number : 11-105767

(22)Date of filing : 13.04.1999

(71)Applicant : CANON INC

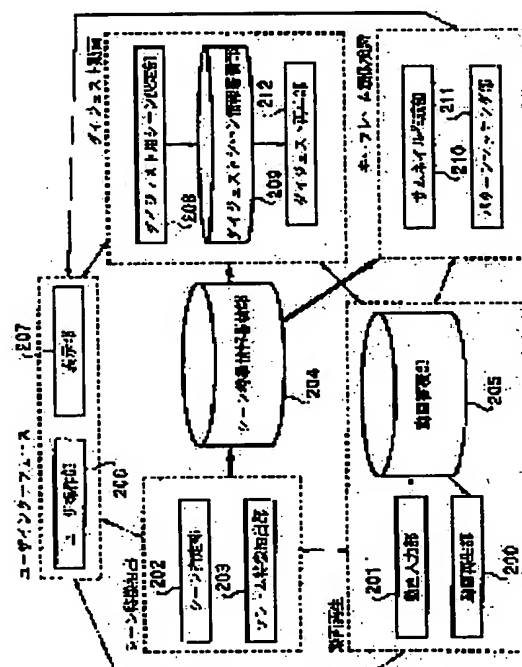
(72)Inventor : IKEDA KAZUYO
SHIYAMA HIROTAKE
TOJO HIROSHI

(54) IMAGE PROCESSING UNIT AND ITS METHOD

(57)Abstract:

PROBLEM TO BE SOLVED: To easily and properly confirm contents of moving picture data.

SOLUTION: A moving picture storage section 205 of this image processing unit stores moving picture data. A scene designation section 202 designates at least one scene in moving picture data and sets a key frame in each designated scene and importance of the designated scene. Furthermore, a frame characteristics extract section 203 acquires a picture characteristics quantity of each designated key frame. A scene characteristics information storage section 204 generates scene characteristics information including a block, a key frame, importance and a picture characteristics quantity of the key frame as to each scene designated by the scene designation section 202 and stores the information to a memory in cross-reference with the moving picture. Digest reproduction to reproduce only a major part of designated moving picture data is executed or desired moving picture data are retrieved by similar picture retrieval of key frames on the basis of the scene characteristics information.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号
特開2000-299829
(P2000-299829A)

(43)公開日 平成12年10月24日 (2000. 10. 24)

(51)Int.Cl. ⁷	識別記号	F I	テ-マ-コ-ト ⁸ (参考)
H 0 4 N 5/76		H 0 4 N 5/76	B 5 B 0 5 0
G 0 6 F 17/30		G 0 6 F 15/40	3 7 0 D 5 B 0 7 5
G 0 6 T 1/00		15/62	P 5 C 0 5 2
H 0 4 N 5/91		H 0 4 N 5/91	N 5 C 0 5 3
5/765			L

審査請求 未請求 請求項の数58 O L (全 20 頁) 最終頁に続く

(21)出願番号 特願平11-105767

(22)出願日 平成11年4月13日 (1999. 4. 13)

(71)出願人 000001007

キヤノン株式会社

東京都大田区下丸子3丁目30番2号

(72)発明者 池田 和世

東京都大田区下丸子3丁目30番2号 キヤ
ノン株式会社内

(72)発明者 椎山 弘隆

東京都大田区下丸子3丁目30番2号 キヤ
ノン株式会社内

(74)代理人 100076428

弁理士 大塚 康徳 (外2名)

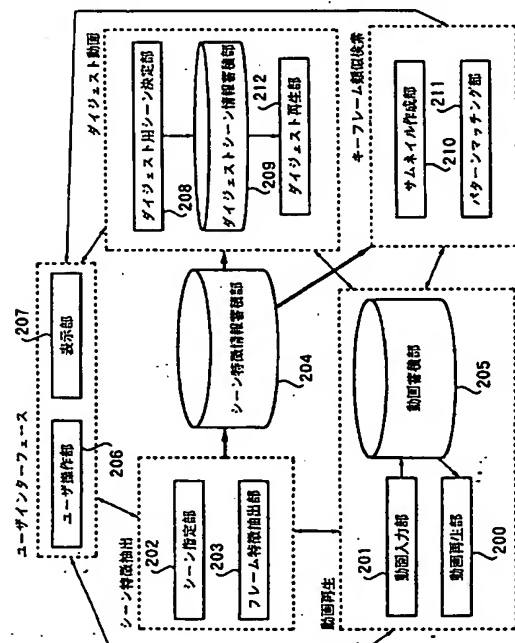
最終頁に続く

(54)【発明の名称】 画像処理装置及び方法

(57)【要約】

【課題】動画像データの内容確認を容易かつ適確に行えるようにする。

【解決手段】動画蓄積部205は動画データを格納する。シーン指定部202は、動画像データ中の少なくとも1つのシーンを指定するとともに、指定された各シーン内におけるキーフレームと、指定されたシーンの重要度を設定する。また、フレーム特徴抽出部203は、指定された各キーフレームの画像特徴量を取得する。シーン特徴情報蓄積部204は、シーン指定部202で指定された各シーンについて、その区間、キーフレーム、重要度、キーフレームの画像特徴量を含むシーン特徴情報を生成し、これを動画像に関連づけてメモリに格納する。このシーン特徴情報に基づいて、指定された動画像データの主要部分のみを再生するダイジェスト再生を実行したり、キーフレームの類似画像検索により所望の動画データを検索する。



【特許請求の範囲】

【請求項 1】 動画像データ中の少なくとも 1 つの所望のシーンと、各所望のシーン内におけるキーフレームとすべきフレームとを特定する特定手段と、

前記特定手段で特定された所望のシーンの各々の重要度を設定する設定手段と、

前記特定手段で特定されたシーンとキーフレームを示す情報、及び前記設定手段で設定された重要度をシーン情報として当該動画像データに関連させて格納する格納手段とを備えることを特徴とする画像処理装置。

【請求項 2】 前記特定手段で特定されたシーンを示す情報は当該シーンの開始フレームと終了フレームのそれぞれのフレーム識別子を含み、キーフレームを示す情報は当該キーフレームのフレーム番号を含むことを特徴とする請求項 1 に記載の画像処理装置。

【請求項 3】 前記キーフレームの特徴量情報を取得する取得手段を更に備え、

前記格納手段において格納されるシーン情報は前記取得手段で取得されたキーフレームの画像の特徴量情報を含むことを特徴とする請求項 1 又は 2 に記載の画像処理装置。

【請求項 4】 前記キーフレームの特徴量情報は、該キーフレームの画像を複数ブロックに分割し、各ブロックについて取得された特徴量に応じてラベルを付与し、付与されたラベルを所定の順序で並べて得られるラベル列であることを特徴とする請求項 3 に記載の画像処理装置。

【請求項 5】 前記格納手段に格納されたシーン情報に基づいて、動画再生が可能な抽出動画データを生成する生成手段を更に備えることを特徴とする請求項 1 又は 2 に記載の画像処理装置。

【請求項 6】 前記生成手段は、前記格納手段に格納されたシーン情報が示すシーンを時系列順にマージすることにより、抽出動画データを生成することを特徴とする請求項 5 に記載の画像処理装置。

【請求項 7】 前記生成手段は、前記格納手段に格納されたシーン情報が示すシーンと、当該動画像の所定時間分の先頭シーン及び終了シーンを時系列順にマージすることにより、抽出動画データを生成することを特徴とする請求項 5 に記載の画像処理装置。

【請求項 8】 前記抽出動画の再生時間を指定する指定手段と、

前記指定手段で指定された再生時間に基づいて、前記生成手段でマージされる各シーンの長さを調整する調整手段とを更に備えることを特徴とする請求項 5 に記載の画像処理装置。

【請求項 9】 前記調整手段は、各シーンに設定された重要度に基づいてシーン内の各フレームの重要度を決定し、該シーン内において閾値を越える重要度を有するフレームを抽出し、これをマージ対

象のシーンとして前記生成手段に提供するフレーム抽出手段と、

前記生成手段で生成される抽出動画データの再生時間が、前記指定された再生時間の所定誤差範囲に納まるように、前記フレーム抽出手段において用いられる閾値を制御する閾値制御手段とを備えることを特徴とする請求項 8 に記載の画像処理装置。

【請求項 10】 前記シーン情報によって指定される全てのシーンの合計時間が前記指定手段で指定された再生時間よりも短い場合、該指定された再生時間の所定誤差範囲に納まるように各シーンのキーフレームを中心として再生時間を暫時延長する延長調整手段を更に備えることを特徴とする請求項 9 に記載の画像処理装置。

【請求項 11】 前記抽出動画データを構成するシーンの中に、その再生時間が所定時間よりも短いものが存在する場合に、これを除去する除去手段を更に備えることを特徴とする請求項 5 に記載の画像処理装置。

【請求項 12】 前記除去手段で除去されたシーンの再生時間に相当する時間分だけ他のシーンを延長する延長手段を更に備えることを特徴とする請求項 11 に記載の画像処理装置。

【請求項 13】 前記延長手段は、前記抽出動画データ内における前記除去手段で除去されたシーンの前後に隣接するシーンのうちの設定された重要度の高いほうのシーンを、キーフレームを中心に時間延長することを特徴とする請求項 12 に記載の画像処理装置。

【請求項 14】 前記抽出動画の再生時間を指定する指定手段と、

前記格納手段に格納されたシーン情報によって示されるシーンをその重要度の順に選択して所定時間長分を切り出す切出し手段と、該切出し手段は、切り出されたシーンの合計時間が前記指定手段で指定された指定時間を越えるまで処理を実行し、

前記生成手段は、前記切出し手段で切り出されたシーンを時系列関係を保ちながら順次マージして抽出動画データを生成することを特徴とする請求項 5 に記載の画像処理装置。

【請求項 15】 前記所定時間長は、ユーザが画像を認識可能な長さであることを特徴とする請求項 14 に記載の画像処理装置。

【請求項 16】 前記シーン情報によって指定される全てのシーンの合計時間が前記指定手段で指定された再生時間よりも短い場合、該指定された再生時間の所定誤差範囲に納まるように各シーンのキーフレームを中心として再生時間を暫時延長する延長調整手段を更に備えることを特徴とする請求項 15 に記載の画像処理装置。

【請求項 17】 前記キーフレームの画像をサムネイル化して一覧表示する表示手段を更に備えることを特徴とする請求項 1 又は 2 に記載の画像処理装置。

【請求項 18】 前記キーフレームの一つが指定された

場合に、該指定されたキーフレームと他のキーフレームの画像の類似度を、各キーフレームの画像特徴量に基づいて算出する算出手段と、

前記算出手段で算出された類似度に基づいてキーフレームを提示する提示手段と、

前記提示手段で提示されたキーフレームの中から選択されたキーフレームを含む動画データを取得し、再生する再生手段とを更に備えることを特徴とする請求項3または4に記載の画像処理装置。

【請求項19】 前記シーン情報は、各キーフレームの画像の画像特徴量を含み、
前記キーフレームの一つが指定された場合に、該選択されたキーフレームと他のキーフレームの画像の類似度を各キーフレームの画像特徴量に基づいて算出する算出手段と、

前記算出手段で算出された類似度に基づいてキーフレームを提示する提示手段と、

前記提示手段で提示されたキーフレームの中から選択されたキーフレームを含む動画データに関して、前記生成手段で生成された抽出画像データを再生する再生手段とを更に備えることを特徴とする請求項5乃至16のいずれかに記載の画像処理装置。

【請求項20】 動画データ中の少なくとも1つの所望のシーンと、各所望のシーン内におけるキーフレームとすべきフレームとを特定する特定工程と、

前記特定工程で特定された所望のシーンの各々の重要度を設定する設定工程と、

前記特定工程で特定されたシーンとキーフレームを示す情報、及び前記設定工程で設定された重要度をシーン情報として当該動画データに関連させて記憶手段に格納する格納工程とを備えることを特徴とする画像処理方法。

【請求項21】 前記特定工程で特定されたシーンを示す情報は当該シーンの開始フレームと終了フレームのそれぞれのフレーム番号を含み、キーフレームを示す情報は当該キーフレームのフレーム番号を含むことを特徴とする請求項20に記載の画像処理方法。

【請求項22】 前記キーフレームの特徴量情報を取得する取得工程を更に備え、

前記格納工程において格納されるシーン情報は前記取得工程で取得されたキーフレームの画像の特徴量情報を含むことを特徴とする請求項20又は21に記載の画像処理方法。

【請求項23】 前記キーフレームの特徴量情報は、該キーフレームの画像を複数ブロックに分割し、各ブロックについて取得された特徴量に応じてラベルを付与し、付与されたラベルを所定の順序で並べて得られるラベル列であることを特徴とする請求項22に記載の画像処理方法。

【請求項24】 前記記憶手段に格納されたシーン情報

に基づいて、動画再生が可能な抽出動画データを生成する生成工程を更に備えることを特徴とする請求項20又は21に記載の画像処理方法。

【請求項25】 前記生成工程は、前記記憶手段に格納されたシーン情報が示すシーンを時系列順にマージすることにより、抽出動画データを生成することを特徴とする請求項24に記載の画像処理方法。

【請求項26】 前記生成工程は、前記記憶手段に格納されたシーン情報が示すシーンと、当該動画の所定時間分の先頭シーン及び終了シーンを時系列順にマージすることにより、抽出動画データを生成することを特徴とする請求項24に記載の画像処理方法。

【請求項27】 前記抽出動画の再生時間を指定する指定工程と、
前記指定工程で指定された再生時間に基づいて、前記生成工程でマージされる各シーンの長さを調整する調整工程とを更に備えることを特徴とする請求項24に記載の画像処理方法。

【請求項28】 前記調整工程は、
各シーンに設定された重要度に基づいてシーン内の各フレームの重要度を決定し、該シーン内において閾値を超える重要度を有するフレームを抽出し、これをマージ対象のシーンとして前記生成工程に提供するフレーム抽出工程と、

前記生成工程で生成される抽出動画データの再生時間が、前記指定された再生時間の所定誤差範囲に納まるように、前記フレーム抽出工程において用いられる閾値を制御する閾値制御工程とを備えることを特徴とする請求項27に記載の画像処理方法。

【請求項29】 前記シーン情報によって指定される全てのシーンの合計時間が前記指定工程で指定された再生時間よりも短い場合、該指定された再生時間の所定誤差範囲に納まるように各シーンのキーフレームを中心として再生時間を暫時延長する延長調整工程を更に備えることを特徴とする請求項28に記載の画像処理方法。

【請求項30】 前記抽出動画データを構成するシーンの中に、その再生時間が所定時間よりも短いものが存在する場合に、これを除去する除去工程を更に備えることを特徴とする請求項24に記載の画像処理方法。

【請求項31】 前記除去工程で除去されたシーンの再生時間に相当する時間分だけ他のシーンを延長する延長工程を更に備えることを特徴とする請求項30に記載の画像処理方法。

【請求項32】 前記延長工程は、前記抽出動画データ内における前記除去工程で除去されたシーンの前後に隣接するシーンのうちの設定された重要度の高いほうのシーンを、キーフレームを中心に時間延長することを特徴とする請求項31に記載の画像処理方法。

【請求項33】 前記抽出動画の再生時間を指定する指定工程と、

前記記憶手段に格納されたシーン情報によって示されるシーンをその重要度の順に選択して所定時間長分を切り出す切出し工程と、該切出し工程は、切り出されたシーンの合計時間が前記指定工程で指定された指定時間を越えるまで処理を実行し、
前記生成工程は、前記切出し工程で切り出されたシーンを時系列関係を保ちながら順次マージして抽出動画データを生成することを特徴とする請求項24に記載の画像処理方法。

【請求項34】 前記所定時間長は、ユーザが動画を認識可能な長さであることを特徴とする請求項33に記載の画像処理方法。

【請求項35】 前記シーン情報によって指定される全てのシーンの合計時間が前記指定工程で指定された再生時間よりも短い場合、該指定された再生時間の所定誤差範囲に納まるように各シーンのキーフレームを中心として再生時間を暫時延長する延長調整工程を更に備えることを特徴とする請求項34に記載の画像処理方法。

【請求項36】 前記キーフレームの画像をサムネイル化して一覧表示する表示工程を更に備えることを特徴とする請求項20又は21に記載の画像処理方法。

【請求項37】 前記キーフレームの一つが指定された場合に、該指定されたキーフレームと他のキーフレームの画像の類似度を、各キーフレームの画像特徴量に基づいて算出する算出工程と、

前記算出工程で算出された類似度に基づいてキーフレームを提示する提示工程と、

前記提示工程で提示されたキーフレームの中から選択されたキーフレームを含む動画像データを取得し、再生する再生工程とを更に備えることを特徴とする請求項22または23に記載の画像処理方法。

【請求項38】 前記シーン情報は、各キーフレームの画像の画像特徴量を含み、
前記キーフレームの一つが指定された場合に、該選択されたキーフレームと他のキーフレームの画像の類似度を、各キーフレームの画像特徴量に基づいて算出する算出工程と、

前記算出工程で算出された類似度に基づいてキーフレームを提示する提示工程と、

前記提示工程で提示されたキーフレームの中から選択されたキーフレームを含む動画像データに関して、前記生成工程で生成された抽出画像データを再生する再生工程とを更に備えることを特徴とする請求項24乃至35のいずれかに記載の画像処理方法。

【請求項39】 コンピュータによる動画像データの格納処理を実現するための制御プログラムを格納する記憶媒体であって、該制御プログラムが、
動画像データ中の少なくとも1つの所望のシーンと、各所望のシーン内におけるキーフレームとすべきフレームとを特定する特定工程のコードと、

前記特定工程で特定された所望のシーンの各々の重要度を設定する設定工程のコードと、

前記特定工程で特定されたシーンとキーフレームを示す情報、及び前記設定工程で設定された重要度をシーン情報として当該動画像データに関連させて記憶手段に格納する格納工程のコードとを備えることを特徴とする記憶媒体。

【請求項40】 前記キーフレームの特徴量情報を取得する取得工程のコードを更に備え、

前記格納工程において格納されるシーン情報は前記取得工程で取得されたキーフレームの画像の特徴量情報を含むことを特徴とする請求項39に記載の記憶媒体。

【請求項41】 前記記憶手段に格納されたシーン情報に基づいて、動画再生が可能な抽出動画データを生成する生成工程のコードを更に備えることを特徴とする請求項39に記載の記憶媒体。

【請求項42】 前記キーフレームの一つが指定された場合に、該指定されたキーフレームと他のキーフレームの画像の類似度を、各キーフレームの画像特徴量に基づいて算出する算出工程のコードと、

前記算出工程で算出された類似度に基づいてキーフレームを提示する提示工程のコードと、

前記提示工程で提示されたキーフレームの中から選択されたキーフレームを含む画像を取得し、再生する再生工程のコードとを更に備えることを特徴とする請求項40に記載の記憶媒体。

【請求項43】 前記シーン情報は、各キーフレームの画像の画像特徴量を含み、

前記キーフレームの一つが指定された場合に、該選択されたキーフレームと他のキーフレームの画像の類似度を各キーフレームの画像特徴量に基づいて算出する算出工程のコードと、

前記算出工程で算出された類似度に基づいてキーフレームを提示する提示工程のコードと、

前記提示工程で提示されたキーフレームの中から選択されたキーフレームを含む動画像データに関して、前記生成工程で生成された抽出画像データを再生する再生工程のコードとを更に備えることを特徴とする請求項41に記載の記憶媒体。

【請求項44】 動画像データを格納する記憶媒体であって、

格納されている動画像データ中より特定された少なくとも1つの所望のシーンを表わすシーン特定情報と、

前記シーン特定情報で特定される各シーンにおけるキーフレームを表わすキーフレーム情報と、

前記シーン特定情報で特定されるシーンの各々について設定された重要度とを含むシーン情報を、当該動画像データに対応づけて格納してあることを特徴とする記憶媒体。

【請求項45】 前記シーン特定情報は、各シーンの開

始フレームと終了フレームのそれぞれのフレーム番号を含み、前記キーフレーム情報は当該キーフレームのフレーム番号を含むことを特徴とする請求項44に記載の記憶媒体。

【請求項46】 前記シーン情報において、前記キーフレーム情報は、各キーフレームの画像の特徴量情報を含むことを特徴とする請求項44に記載の記憶媒体。

【請求項47】 前記複数ブロックへの分割は、画像を縦横共に2のべき乗でブロックに分割するものであることを特徴とする請求項44に記載の画像処理装置。

【請求項48】 前記ラベル列とともに画像の分割数、特徴量の抽出方法、および色表現に関する情報を記憶する手段を備えることを特徴とする請求項4又は47に記載の画像処理装置。

【請求項49】 前記シーン情報は、各キーフレームの画像の画像特徴量を含み、
前記キーフレームをランダムに表示する第1手段と、
前記生成手段で生成された抽出画像データを再生する第2手段と、
指定されたキーフレームと他のキーフレームの画像の類似度を各キーフレームの画像特徴量に基づいて算出し、類似検索を行う第3手段と、
指定されたキーフレーム～動画を再生する第4手段とを備え、
前記第1乃至第4手段を任意の組み合わせで機能させるべく操作可能としたことを特徴とする請求項5乃至16のいずれかに記載の画像処理装置。

【請求項50】 前記フレーム識別子は、動画データにおける先頭からのフレーム番号であることを特徴とする請求項2に記載の画像処理装置。

【請求項51】 前記フレーム識別子は、動画データにおける先頭から当該フレームまでの時間であることを特徴とする請求項2に記載の画像処理装置。

【請求項52】 前記キーフレームの特徴量情報は、該キーフレームの画像を複数ブロックに分割し、各ブロックについて取得された特徴量に応じてラベルを付与し、付与されたラベルを所定の順序で並べて得られるラベル列であることを特徴とする請求項19に記載の画像処理装置。

【請求項53】 前記複数ブロックへの分割は、画像を縦横共に2のべき乗でブロックに分割するものであることを特徴とする請求項23に記載の画像処理方法。

【請求項54】 前記ラベル列とともに画像の分割数、特徴量の抽出方法、および色表現に関する情報を記憶する工程を備えることを特徴とする請求項23又は53に記載の画像処理方法。

【請求項55】 前記シーン情報は、各キーフレームの画像の画像特徴量を含み、
前記キーフレームをランダムに表示する第1工程と、
前記生成工程で生成された抽出画像データを再生する第

2工程と、

指定されたキーフレームと他のキーフレームの画像の類似度を各キーフレームの画像特徴量に基づいて算出し、類似検索を行う第3工程と、

指定されたキーフレーム～動画を再生する第4工程とを備え、

前記第1乃至第4工程を任意の組み合わせで機能させるべく操作可能としたことを特徴とする請求項24乃至35のいずれかに記載の画像処理方法。

10 【請求項56】 前記フレーム識別子は、動画データにおける先頭からのフレーム番号であることを特徴とする請求項21に記載の画像処理方法。

【請求項57】 前記フレーム識別子は、動画データにおける先頭から当該フレームまでの時間であることを特徴とする請求項21に記載の画像処理方法。

【請求項58】 前記キーフレームの特徴量情報は、該キーフレームの画像を複数ブロックに分割し、各ブロックについて取得された特徴量に応じてラベルを付与し、付与されたラベルを所定の順序で並べて得られるラベル列であることを特徴とする請求項38に記載の画像処理方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、動画データの検索操作を容易に行うのに適した形態にて動画データを格納するための画像処理装置及び方法に関する。

【0002】

【従来技術】 近年、MPEG等の動画データ圧縮技術の進歩と、記録メディアの大容量化により、インターネット等を用いて動画データを配信することが実現されている。また、個人ユーザレベルにおいても、大量のデジタル画像データを格納しておくことが一般に行える状況となっている。

【0003】 このような大量の動画データの中から自分の見たい動画を選択するには、操作者が一つ一つの動画データについて倍速再生を行ったり、早送りと再生を繰り返したりすることにより、その内容を確認する必要があった。

【0004】

40 【発明が解決しようとする課題】 しかしながら、従来のような方法では、ユーザは面倒な操作を強いられ、所望の動画にたどり着くまでに大変な労力と時間を費やさなければならない。

【0005】 本発明は上記の問題に鑑みてなされたものであり、動画データの内容確認を容易かつ適確に行える画像処理装置及び方法を提供することを目的とする。

【0006】 また、本発明の他の目的は、上記目的を達成するのに好適なデータ構造を提供することにある。

【0007】 また、本発明の他の目的は、所望の動画データを得るための検索を容易に行えるようにすること

にある。

【0008】

【課題を解決するための手段】上記の目的を達成するための本発明の画像処理装置は例えば以下の構成を備える。すなわち、動画像データ中の少なくとも1つの所望のシーンと、各所望のシーン内におけるキーフレームとすべきフレームとを特定する特定手段と、前記特定手段で特定された所望のシーンの各々の重要度を設定する設定手段と、前記特定手段で特定されたシーンとキーフレームを示す情報、及び前記設定手段で設定された重要度をシーン情報として当該動画像データに関連させて格納する格納手段とを備える。

【0009】また、好ましくは、前記キーフレームの特徴量情報を取得する取得手段を更に備え、前記格納手段において格納されるシーン情報は前記取得手段で取得されたキーフレームの画像の特徴量情報を含み、前記キーフレームの一つが指定された場合に、該指定されたキーフレームと他のキーフレームの画像の類似度を、各キーフレームの画像特徴量に基づいて算出する算出手段と、前記算出手段で算出された類似度に基づいてキーフレームを提示する提示手段と、前記提示手段で提示されたキーフレームの中から選択されたキーフレームを含む画像を取得し、再生する再生手段とを備える。

【0010】また、好ましくは、前記シーン情報は、各キーフレームの画像の画像特徴量を含み、前記格納手段に格納されたシーン情報に基づいて、動画再生が可能な抽出動画データを生成する生成手段と、前記キーフレームの一つが指定された場合に、該選択されたキーフレームと他のキーフレームの画像の類似度を各キーフレームの画像特徴量に基づいて算出する算出手段と、前記算出手段で算出された類似度に基づいてキーフレームを提示する提示手段と、前記提示手段で提示されたキーフレームの中から選択されたキーフレームを含む画像に関して、前記生成手段で生成された抽出画像データを再生する再生手段とを備える。

【0011】

【発明の実施の形態】以下、添付の図面を参照して本発明の好適な実施形態を説明する。

【0012】＜第1の実施形態＞図1は本実施形態の画像処理装置の制御構成を示すブロック図である。図1において、101はCPUであり、本実施形態の画像処理装置における各種制御を実行する。102はROMであり、本装置の立ち上げ時に実行されるブートプログラムや各種データを格納する。103はRAMであり、CPU101が処理するための制御プログラムを格納するとともに、CPU101が各種制御を実行する際の作業領域を提供する。104はキーボード、105はマウスであり、ユーザによる各種入力操作環境を提供する。

【0013】106は外部記憶装置であり、ハードディスクやフロッピーディスク、CD-ROM等で構成され

る。107は表示器であり、CRTディスプレイや液晶ディスプレイなどで構成され、CPU101の制御の下で各種表示を行う。108はネットワークインターフェースであり、ネットワーク上の各機器との通信を可能とする。109はビデオインターフェースであり、ビデオカメラ110や、VTR112からの動画データの取り込みを可能とする。また、111は、上記の各構成を接続するバスである。

【0014】なお、上記の構成においてビデオカメラ110、VTR112や外部記憶装置106はネットワーク上に配置されたもので代用してもよい。

【0015】図2は本実施形態の画像処理装置の機能構成を示すブロック図である。図2において、201は動画入力部であり、ビデオカメラ110、VTR112等からビデオインターフェース109を介して画像を取り込む、或いは、ネットワーク上からネットワークインターフェース108を介して動画を取り込む。205は動画蓄積部であり、動画入力部201で取り込まれた動画の全フレームを蓄積するものである。200は動画再生部であり、指定されたフレーム番号に基づいて動画蓄積部205から動画フレームを順次呼び出すことで動画を再生する。

【0016】207は表示部であり、動画再生部200によって再生される動画や、サムネイル作成部210によって作成されるサムネイル画像を、ユーザに対して表示器107を介して表示する。ユーザ操作部206は、ユーザがキーボード104、マウス105を用いて、動画再生の指示、ランダム表示指示、類似検索指示、ダイジェスト再生指示、等を行うものである。

【0017】202はシーン指定部であり、重要なシーンの開始フレーム、終了フレーム、及び、キーフレームの指定を行う。203はフレーム特徴抽出部であり、シーン指定部202で指定されたキーフレームからカラーレイアウトを抽出する。なお、本実施形態においては、カラーレイアウトは特開平10-260983に開示されているように、フレーム画像を分割してえられる複数のブロックの特徴量をラベル化し、これを所定のルールで並べたラベル列が用いられる。204はシーン特徴情報蓄積部であり、シーン指定部202で指定されたデータとフレーム特徴抽出部203で抽出されたカラーレイアウトをひとつのデータ構造にまとめて蓄積する。

【0018】208はダイジェスト用シーン決定部であり、シーン特徴情報蓄積部204を参照してダイジェスト用シーンを決定するものである。209はダイジェストシーン情報蓄積部であり、ダイジェスト用シーン決定部208によって決定されたダイジェスト用シーンの情報を格納する。212はダイジェスト再生部であり、ダイジェストシーン情報蓄積部209を参照し、動画再生部200を介して、ダイジェストの再生を行う。

【0019】210はサムネイル作成部であり、シーン

指定部202により指定されたキーフレーム画像を縮小してサムネイル画像を作成する。211はパターンマッチング部であり、キーフレームのカラーレイアウトのラベル列同士のマッチングを行うものである。

【0020】図3は本実施形態の画像処理装置に格納されるシーン特徴データのデータ構造を示す図である。シーン特徴情報蓄積部204は、図3の(a)に示されるごときシーン特徴データを、当該シーン特徴データが抽出された動画データに対応づけて格納する。シーン特徴データにおいて、開始フレーム番号は、当該シーンの10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

【0021】図3の(b)はキーフレームのカラーレイアウトのデータ構造を示す図である。また、図4は本実施形態による画像のブロック分割例を示す図である。以下、図3及び図4を参照して、本実施形態における「キーフレームのカラーレイアウト」について更に詳細な説明を加える。

【0022】図3の(b)において、「分割数」は画像の分割数である。本実施形態では、画像の分割を縦横同じ分割数でかつ2のべき乗の数とし、そのべきの数で分割数を表現する。例えば、図4に示すように、縦横の分割数がそれぞれ8であれば、 $8 = 2^3$ であるから、3で表されることになる。

【0023】また、「色特徴の抽出方法」は、画像を分割して得られた各ブロックにおける特徴量の抽出方法を示す。例えば、平均色(画素全体の色の平均)、最頻色(画素全体の中で最も頻度の高い色)を用いることなどが上げられる。

【0024】「カラーデータフォーマット」は、色データの形式を示すものである。図3の(c)に、カラーデータフォーマットのデータ構造を示す。「色空間」は色を表現するための空間であり、例えばRGB空間や、 $L^*u^*v^*$ 空間などである。「色の深さ」は、1画素あたり何bitで表現するかであり、32bit/pixel, 24bit/pixel, 16bit/pixelなどである。「カラーマップ」は色番号と実際の色を対応付けるマップである。ただし、「カラーマップ」は、必ずしも必要ではない。

【0025】「ラベル列」は、各ブロックについて取得された特徴量に応じて付与されたラベルの列である。この数はもちろん、ブロックの分割数によって決まる。以下、ラベル列の生成について説明する。

【0026】図5は本実施形態による多次元特徴量空間を説明する図である。図5に示すように、多次元特徴量空間(RGBカラー空間)を複数の色セルに分割し、夫々の色セルに対して通し番号でユニークなラベルを付与する。ここで、多次元特徴量空間(RGBカラー空間)を複数の色セルに分けたのは微妙な特徴量(色)の違いを吸収するためである。なお、多次元特徴量空間に関しては、画像特徴量をそのまま用いるのではなく各パラメータを平均と分散を実験によって求め規格化(正規化)した後、例えば、主成分分析等の直交変換を行い、意味のある次元にしたものを用いることが考えられる。なお、「意味のある次元」とは、主成分分析において、寄与率が大きな主成分軸で構成される次元である。

【0027】次に、キーフレーム画像を分割して得られた各ブロックに対して、「色特徴の抽出方法」によって定められた画像特徴量計算処理を行い、上記多次元特徴量空間上のどのセルに属するかを求め、対応するラベルを求める。この処理を全てのブロックに対して行う。すなわち、各ブロックにおいて全ての画素がどの色セルに属するか計算処理を行い、もっとも頻度の多い色セルのラベルをその分割画像ブロックのラベルとして決定し、この処理を全てのブロックに対して行う。

【0028】以上のようにして各ブロックに対してパラメータラベルが付与されると、各ブロックに付与されたパラメータラベルを所定のブロック順序で並べることにより、ラベル列が生成される。例えば、図4に示されるように、左から右への走査を上から下へ行うことによりパラメータラベルを並べ、ラベル列を作る。

【0029】以上のような構成を備えた本実施形態の画像処理装置の動作例を以下に説明する。

【0030】[シーン特徴抽出処理] まず、動画から重要なシーンに関する情報を抽出しシーン特徴データ(図3)として記憶する処理を説明する。

【0031】図6は本実施形態による登録処理の手順を表すフローチャートである。まず、ステップS11において、シーン指定部202は、動画中の重要なシーンの開始フレーム、終了フレーム、当該シーンのキーとなるキーフレームを決定するとともに、当該シーンの重要度を決定する。この処理は、人手によってマニュアルで行っても、何らかの演算処理を用いて自動的に行っても構わない。以下では、マニュアルによって決定を行う例を示す。

【0032】まず、動画再生部200は、動画入力部201によって入力され動画蓄積部205に記憶された動画を読み出す。ユーザによるユーザ操作部206を介した指示によって再生、早送り、巻き戻し等を行う。そして、ユーザは、表示部207に再生される画像を観察し、重要なシーンを指定する。例えば、画像を再生させながら、重要シーンの開始場面(フレーム)と終了場面(フレーム)を指示するとともに、当該重要シーンの中

で最もよくそのシーンを表わしている場面（フレーム）をキーフレームとして指示する。

【0033】シーン指定部202では、ユーザ操作部206を介して指示された開始フレームと、終了フレームと、キーフレームのそれぞれのフレーム番号を得る。さらに、当該動画全体に対する当該シーンの重要度を、例えば100を最大値とした整数を用いて指定する。こうして得られたそれぞれの値は、シーン指定部202によって図3のシーン特徴データの形式にまとめられる。

【0034】次に、こうして得られたシーン特徴データは、ステップS12において、対象としている動画の1Dと当該動画から抽出されたシーンの先頭から割り振られたシーン1Dと共にシーン特徴情報としてシーン特徴情報蓄積部204に格納される。図7は、シーン特徴情報の一例を示す図である。

【0035】次にステップS13において、フレーム特徴抽出部203は、シーン特徴情報蓄積部204を参照して、動画蓄積部205よりキーフレームを順次取り出し、色に関する特徴抽出を行って「キーフレームのカラーレイアウト」を決定し、シーン特徴情報蓄積部204に格納していく。カラーレイアウトを決定する処理については上述したとおりであり、更に詳細な説明は、特開平10-260983に述べられている。この処理はシーン特徴情報に含まれる全てのキーフレーム番号について行われる。

【0036】以上のようにして、ユーザによって指定された全てのシーンについて、シーン特徴データ（図3）が抽出され、シーン特徴情報としてシーン特徴情報蓄積部204に格納される。以上がシーン特徴抽出の処理であり、この処理を対象となる全ての動画に対して行う。

【0037】【ダイジェスト動画作成処理】次に、シーン特徴情報からそれぞれの動画についてのダイジェスト動画を自動作成する処理を説明する。図8及び図9は本実施形態による（一つの動画に対する）ダイジェスト動画作成処理の手順を表すフローチャートである。

【0038】まず、ステップS21において、ダイジェスト用シーン決定部208は、シーン特徴情報蓄積部204を参照し、同一の動画1Dをもつ全てのシーン特徴データの開始フレーム番号と終了フレーム番号から、その合計時間を算出する。この合計時間と、ユーザによって指定されたダイジェスト動画再生のための指定時間から動画の先頭及び終了シーンの時間を差し引いた時間

（以下、これをTSと表わす）とを比較する。ここで指定時間から先頭及び終了シーンの時間を除くのは、ダイジェスト動画には動画像の先頭と終了のシーンが必ず含まれるという観点に基づいている。また、先頭及び終了シーンの時間は、人間の視覚特性に基づき目にちたつかない程度の最適な長さのものとする。よってこの時間長に関しては特に限定しないが、数秒オーダーのものである。なお、先頭と終了のシーンを必ずしも含めないよう

に構成することももちろん可能である。

【0039】さて、ステップS21において、シーンの合計時間の方が長い場合は、シーンの合計時間を指定時間に合わせるべくステップS22以降の処理を実行する。

【0040】まず、ステップS22では、全てのシーン特徴データの重要度からその最大値を取得する。次に、ステップS23において、重要度の下限の閾値と最大値の中間値を、重要度の閾値とする。ここで、重要度の下限の閾値とは、人間が重要なシーンであると判定する際の最低限の重要度である。すなわち、重要なシーンと判定されたシーンの開始フレームと終了フレームが、この重要度を持っているものと仮定される。なお、この重要度の下限の閾値は経験的に決まるものであり、特に限定されるものではない。次に、ステップS24において、後続のステップS25からS32までの一連の処理のために、初期値としてSaに重要度の最大値を、Sbに重要度の下限の閾値を設定する。

【0041】次に、ステップS25において、閾値以上の重要度を持つダイジェスト用シーンを決定する。この処理について、図10を用いてより詳細に説明する。図10には、動画とその重要なシーンを概念的に表したものと、これに対応させて、横軸に時間、縦軸にフレームの重要度をとったグラフが示されている。先述したようにそれぞれの重要なシーンの開始フレーム、終了フレームの重要度は閾値の下限値と仮定される。また、キーフレームはそのシーンを最もよく表しているものであるから、シーンの重要度がキーフレームの重要度に相当するものとする。なお、重要なシーン以外のフレームの重要度は不定であるとする（点線で図示された部分）。

【0042】以上から、開始フレームの重要度からキーフレームの重要度までの間と、キーフレームの重要度から終了フレームの重要度までの間に仮想的な直線を引くことができる（図10において実線で示された部分）。これらの仮想的な重要度の直線と閾値の交点を求め、その交点の間に存在するフレームが、閾値以上の重要度をもったフレームとして抽出され、これら抽出されたフレームによってダイジェスト用シーンが決定される。すなわち、図10では（a）から（b）までのフレームと（c）から（d）までのフレームがこれに相当し、これらのフレームがダイジェスト用シーンとなる。

【0043】次にステップS26において、ステップS25で決定されたダイジェスト用シーンの合計時間からTSを減算し、これをTRとする。次に、ステップS27において、TRが0以上であり、かつ誤差の時間以内かどうかを判定する。ここで誤差の時間とは、閾値をどの位置に動かしても、TRが0とならない（即ち、ダイジェスト用のシーンの合計時間がTSに一致しない）場合を想定して設けられた時間であり、経験的に求まるものである。ここで、TRが所定の時間以上であれば、次

にステップS28に進む。

【0044】ステップS28では、TRが正であるかを判定する。ここで正であると判定された場合は、ダイジェスト用シーンの合計時間が長すぎることを意味するので、ステップS29へ進み、現在の閾値をSbに設定する。そして、ステップS30において、閾値をSaと現在の閾値の中間値に更新する。そして、ステップS25へ戻る。すなわち、重要度の下限の閾値を現在設定されている閾値に設定し、この新たに設定された下限の閾値とSaに格納されている最大値との中間値を新たに閾値に設定して、上述のステップS25以降の処理を繰り返す。

【0045】一方、ステップS27において負であると判定された場合は、ダイジェスト用シーンの合計時間が短すぎることを意味し、処理はステップS31へ進む。ステップS31においては、現在の閾値をSaに設定する。そして、ステップS32において、閾値をSbと現在の閾値の中間値に更新する。そして、ステップS25へと戻る。すなわち、重要度の最大値を現在設定されている閾値に設定し、この新たに設定された最大値とSbに格納されている下限値との間の中間値を新たな閾値に設定として、上述のステップS25以降の処理を繰り返す。

【0046】以上のようにステップS25からステップS32を繰り返すことで、ダイジェスト用シーンは、TSに最も近くなり、最終的にステップS27で、TRは0以上で所定の時間より短いと判定され、ステップS34へと進むことになる。

【0047】一方、ステップS21で、シーンの合計時間の方が短いと判定された場合は、このままでは、ダイジェスト動画の時間がユーザの指定した時間に満たないので、ステップS33へ進み、例外処理を行う。ステップS33では、シーン特徴情報に登録されている当該動画画像の全ての重要なシーンについて、そのキーフレームを中心として、それぞれ暫時、シーンを長くしていき、これら重要シーンの再生時間をTSに近づける。すなわち、キーフレームの左側と右側交互にフレームを追加していき、シーンを長くする。そして、ステップS34へ進む。

【0048】ステップS34において、ダイジェスト用シーンとして抽出、決定されたシーンは時系列にソートされ、図11に例示するようなダイジェストシーン情報として、ダイジェストシーン情報蓄積部209へ蓄積される。次に、ステップS35において、当該動画画像の先頭シーンをダイジェストシーン情報の先頭に、終了シーンをダイジェストシーン情報の最後に格納する。もしこのとき、前後のシーンと重要シーンのフレームが重なるときは、マージを行う。

【0049】次に、ステップS36において、ダイジェストシーン情報の各シーンを順に調べて、短すぎるシー

ンが存在しないかどうかチェックを行う。ここで短すぎるシーンとは、人間が目でその内容を判定不可能と思われる時間であり、人間の視覚特性に基づき決定されるものである。本例では、先頭及び終了シーンの時間として定めたものより短いものとする。ここで短すぎるシーンが存在した場合は、ステップS37において、その短すぎるシーンをダイジェストシーン情報より取り除く。次に、ステップS38において、ステップS37で取り除いたシーンの前後にあるシーンのうち、その重要度の高い方のシーンについて、キーフレームを中心にシーンを暫時長くしていき（ステップS33と同様）、取り除いた分を埋める。次に、ステップS39において、ダイジェストシーン情報中の全てのシーンについてチェックを行ったかどうかを判定し、全てについてチェックが行われるまでステップS36からステップS38を繰り返す。

【0050】以上の処理によって得られたダイジェストシーン情報から、それぞれのシーンを順に再生させることにより、動画データに対するダイジェスト動画を得ることができる。以上が、ダイジェスト動画作成処理であり、この処理を対象となる全ての動画に対して行う。

【0051】【動画検索処理】次に、ユーザが所望の動画を検索する際の処理について述べる。図12は本実施形態による動画検索の際のユーザの典型的な操作における状態遷移図である。また、図13は、本実施形態による動画検索の処理を説明するフローチャートである。

【0052】まず、ユーザのユーザ操作部206を介してランダム表示が指示されると、処理はステップS51からステップS52へ進み、状態：キーフレームランダム表示100となる。このとき、サムネイル作成部210は、シーン特徴情報蓄積部204によって格納されたシーン特徴情報を参照して所定数のシーンIDに含まれるキーフレームの番号を取り出し（ステップS52）、これを基に動画蓄積部205よりキーフレーム画像を取り出す（ステップS53）。そして、キーフレーム画像を縮小してこれをサムネイルとして、ランダムに表示部207を介して表示する（ステップS54）。なお、ステップS52における所定数とは、一画面で表示するサムネイル画像の数に対応する。この場合、これらのサムネイル画像は、動画ID、シーンID、キーフレーム番号と関連づけて特に図示しないメモリ上に保持してもよいし、蓄積媒体上に記憶しておいてもよい。このようにサムネイルを保持しておけば、ステップS54でその都度サムネイルを生成する必要がなくなる。もちろん、検索処理の初期化として、全てのキーフレームのサムネイル画像を作成してこれをメモリ又は蓄積媒体上に動画IDなどと関連づけて保持しておき、ランダム表示の指示があるごとに、所定数のサムネイル画像をメモリ又は蓄積媒体より取り出して表示するようにしてもよい。

【0053】ユーザはキーフレームのサムネイル画像の

中から、所望の動画中に含まれるシーンに似ているものが見つかるまで、このランダム表示の指示を繰り返す（ステップS55）。そして、似ているキーフレーム（所望のキーフレーム）が見つかったら、ユーザ操作部206を介して所望のキーフレームを指定することにより類似フレームの表示を指示する（ステップS56）。この指示により処理はステップS57以降へ進み、画像処理装置は、状態：類似フレーム表示101となる。このとき、パターンマッチング部211は、指定されたキーフレームを比較元として、シーン特徴情報蓄積部204に含まれるキーフレームのカラーレイアウトに含まれるラベル列同士でマッチングを行い、各動画像中のキーフレームとの類似度を算出する（ステップS57）。

【0054】この類似度の算出には、例えば特開平10-260983号に記載されたラベル列間のマッチング処理を適用することが可能である。簡単に説明すると次のとおりである。まず、図14に示すようなペナルティマトリクスを用意する。このペナルティマトリクスでは、図5で示した特徴量空間の各色セルにおいて、隣接する色セル間には小さいペナルティ（距離）を与え、遠い色セル間には大きいペナルティ（距離）を与えるようにラベル間のペナルティが設定されている。そして、比較元のキーフレームのラベル列と他のキーフレームのラベル列との距離を、このペナルティマトリクスを参照して算出し、類似度を得る。特に、曖昧なパターンマッチングを行えるように、音声認識等において用いられるDPマッチング手法や、特開平8-241335号に開示されているようなファジー非決定性有限オートマトンを適用することも可能である。

【0055】また、このとき比較元と比較先においてキーフレームのイメージの「分割数」などの属性情報が異なるときは、それぞれの属性情報を元にラベル列のデータ変換を行うことが可能である。例えば、比較元の分割数が3（ $2^3 \times 2^3 = 64$ 個のブロックとなる）であり、比較先の分割数が2（ $2^2 \times 2^2 = 16$ 個のブロックとなる）であり、色特徴の抽出方法が共に「平均色」である場合は、比較元のラベル列を、4つ一組としてそれぞれ一旦対応する色セルの重心にあたる中間色に直し、この4つの中間色の平均を求め、この平均色に対応するラベルを新たなラベルとしていくことにより、比較先と同じ16個のラベルを含むラベル列に変換できる。

【0056】以上のようにして各キーフレームとの類似度が求まると、類似度の高い順番に表示部207を介して対応するサムネイル画像を検索結果として表示する（ステップS58、S59）。

【0057】ユーザは所望するシーンと思われるキーフレームが得られるまで、これらランダム表示、及び、類似フレーム表示を繰り返す（ステップS60）。次に、ユーザがキーフレーム類似検索結果の中から所望の動画中のシーンと思われるキーフレームのサムネイル画像を

ユーザ操作部206を介して指定し、再生の指示を出すと画像処理装置は、状態：動画再生102となる（ステップS61）。このとき、動画再生部200は、指示されたキーフレームに関連付けられた動画IDとキーフレーム番号を参照し、動画蓄積部205から対応する動画のキーフレーム番号から順次フレーム画像を取り出して表示部207を介して表示する（ステップS63）。

【0058】なお、再生させたシーンがユーザの所望する動画に含まれるシーンでなければ、ランダム表示を再度指示することにより（ステップS60）、ランダム表示を再度実行させたり、他のサムネイル画像を比較元に指定しなおして類似シーン表示を行わせることができる。

【0059】そして、所望の動画に含まれるシーンであると思われる、動画全体を確認するためにユーザはユーザ操作部206を介してダイジェスト再生の指示を出す。ダイジェスト再生が指示されると（ステップS62）、画像処理装置は、状態：ダイジェスト再生103となる（ステップS64）。このときダイジェスト再生部212は、ダイジェスト蓄積部209のダイジェストシーン情報を参照して、指定されたキーフレームの動画IDにおける全てのダイジェスト用シーンの開始フレーム番号から終了フレーム番号を、動画再生部200に通知する。動画再生部200は、通知された開始フレーム番号から終了フレーム番号までのフレームを動画蓄積部から順次取り出して、表示部207を介して表示する。このようにして動画のダイジェストが表示され、ユーザは所望の動画であったかどうかの判定が行える。

【0060】そして、もし所望の動画でなければ、ランダム表示や、他のサムネイル画像を比較元として類似シーン表示を行う（ステップS60）。以上のことを繰り返すことによって、ユーザは所望の動画を得ることが可能となる。

【0061】もちろん、上記の処理の流れは一例であり、それぞれの操作を自由に組み合わせることで、所望の動画に辿り着く事も可能である。例えば、状態：類似フレーム表示101から、ダイジェスト再生指示を与えて、状態：ダイジェスト再生103に移り、動画全体のチェックを行っても構わない。或いは、状態：キーフレームランダム表示100において所望のキーフレームを選択して動画再生やダイジェスト再生を行わせるように構成することもできる。

【0062】以上説明したように、上記第1の実施形態によれば、シーン特徴というデータ構造で動画の重要なシーンに関する情報が記述できるため、キーフレームのランダム表示、キーフレーム類似検索、及び、ダイジェスト自動作成とダイジェスト再生が可能となり、ユーザはこれらを好みの組み合わせで操作することによって、瞬時に所望の動画にたどり着くことが可能となる。また、シーン特徴というデータ構造に従えば、他の画像処

理装置で利用されている動画像も同様に扱うことが可能となる。

【0063】＜第2の実施形態＞上記第1の実施形態においては、ダイジェスト動画作成処理において、指定の時間に納まるように閾値を変化させて、ダイジェスト用シーンを決定している。しかしながら、指定時間に収めたダイジェスト用シーンの決定方法はこれに限られるものではない。例えば、この処理の他の実施形態として図15に示した処理手順を用いることも可能である。

【0064】図15は、第2の実施形態によるダイジェスト用シーンの決定処理手順を説明するフローチャートである。まず、ステップS71において、ダイジェスト用シーン決定部208は、シーン特徴情報蓄積部204を参照し、同一の動画IDをもつ全てのシーン特徴の開始フレーム番号と終了フレーム番号を得て、これに基づいて当該動画の重要シーンとして指定されたシーンの合計時間を算出する。そして、この合計時間と、ユーザによって指定されたダイジェスト動画の再生時間（指定時間）から開始シーンと終了シーンとを差し引いた時間TSとを比較する。ここでシーンの合計時間の方がTSよりも長いときは、このままではダイジェスト再生時間が指定時間を越えてしまうので、ステップS72へ進み、再生時間の調整を行う。

【0065】まず、ステップS72において、シーン特徴情報の同一の動画IDを持つシーン特徴データについて、シーンの重要度でソートを行う。次にステップS73において、この重要度の高い順番に、所定の時間の長さでシーンを取り出し、これをダイジェスト用シーンとする。このとき、所定の時間とは人間の目にちらつかない程度の長さであり、人間の視覚特性によって決定される。また、各シーンから所定の時間分を取り出すときは、キーフレームを中心にして所定時間分のフレームを取り出す。

【0066】次にステップS74において、ステップS73の処理を最も重要度の低いシーンまで行ったか、即ち、ソートされた重要シーンの全てについて行ったかを判定する。ここで、全てのシーン特徴情報についてステップS73の処理が行われたということは、取得したシーンによるダイジェスト再生時間が指定時間以下であることを意味するので、ステップS78へ進み各シーンの再生時間を長くする方向で調整を行う。

【0067】一方、ステップS74において、ステップS73の処理が全ての重要シーンについて行われていないと判断されると、処理はステップS75へ進む。ステップS75では、ダイジェスト用シーンの合計時間がTS以上となったかどうかを判定し、TS以上となるまでステップS73（及びS74）を繰り返す。そして、シーンの合計時間がTS以上となった時点でステップS76へ処理を移す。この結果、ダイジェスト用シーンに重要度の高いシーンが優先的に含まれることになる。

【0068】ここでステップS71においてシーンの合計時間の方が短いと判定された場合、及び、ステップS74において最も重要度の低いシーンまで取得されたと判定された場合の処理であるステップS78について説明する。ステップS78へ処理が移行してきた場合、このままでは、ダイジェスト動画の再生時間がユーザの指定した時間に満たない。従って、ステップS78では、この時点でダイジェスト用シーンを構成する全てのシーンについて、そのキーフレームを中心としてそれぞれ暫時シーンを長くしていき、シーンの合計時間をTSに近づけていく。

【0069】ステップS76においては、ダイジェスト用シーンとして抽出されたシーンを時系列にソートし、図11に例を示したダイジェストシーン情報として、ダイジェストシーン情報蓄積部209へ蓄積する。次に、ステップS77において、先頭のシーンをダイジェストシーン情報の先頭に、終了のシーンをダイジェストシーン情報の最後に格納する。もしこのとき、前後のシーンとフレームが重なるときは、マージを行う。

【0070】以上のような処理によってもダイジェスト動画を自動的に得ることが可能である。

【0071】【他の実施形態】上記各実施形態においては、ダイジェストの作成は、ダイジェストシーン情報を対象となる全ての動画の対して求めた後、一旦、蓄積媒体上に記憶している。しかしながら、ダイジェスト再生指示によりその都度、指示のあった動画のみについてメモリ上にダイジェストシーン情報を作成するようにしてもよい。この場合、本実施形態で説明したダイジェスト動画作成処理そのものは軽いので、即座にダイジェスト再生を行うことが可能である。

【0072】また、上記各実施形態においては、シーン特徴抽出処理と動画検索処理（及びダイジェスト動画作成処理）を一つの画像処理装置で実現しているが、シーン特徴のデータ構造を持ったデータと原動画像及びこれらに関連付ける仕組みが存在すれば、それぞれを別々の画像処理装置で実現しても構わない。

【0073】また、上記各実施形態においては、各フレームの位置を示すのに動画の先頭から割り振られるフレームの番号（開始フレーム番号等）で表現したがこれに限られるものではない。例えば、先頭から再生したときの時間であってもよく、動画中でそのフレームを特定できれば何であってもよい。

【0074】また、上記各実施形態においては、シーン特徴抽出処理において、開始フレーム、終了フレーム、キーフレームの決定を手動で行う場合を説明したが、自動的に各フレームを決定することも可能である。例えば、動画中の全てのフレームに対して特徴抽出を行ってカラーレイアウトを求め、近傍のフレーム間でラベル列の比較を行ってその類似度を算出する。類似度が低いほどフレーム間の変化が激しく、この変化が激しい部分が

心理的に印象に残る。このことを利用して類似度から各フレームの重要度を決定するといった方法が挙げられる。

【0075】なお、本発明は、例えばホストコンピュータ、インターフェース機器、リーダ、プリンタなどの複数の機器から構成されるシステムに適用しても、一つの機器からなる装置（例えば、複写機、ファクシミリ装置など）に適用してもよい。

【0076】また、本発明の目的は、前述した実施形態の機能を実現するソフトウェアのプログラムコードを記憶した記憶媒体を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ（またはCPUやMPU）が記憶媒体に格納されたプログラムコードを読み出し実行することによっても、達成されることは言うまでもない。

【0077】この場合、記憶媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。

【0078】プログラムコードを供給するための記憶媒体としては、例えば、フロッピーディスク、ハードディスク、光ディスク、光磁気ディスク、CD-ROM、CD-R、磁気テープ、不揮発性のメモ리카ード、ROMなどを用いることができる。

【0079】また、コンピュータが読み出したプログラムを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼動しているOS（オペレーティングシステム）などが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0080】さらに、記憶媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書き込まれた後、そのプログラムコードの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現さ

れる場合も含まれることは言うまでもない。

【0081】

【発明の効果】以上説明したように、本発明によれば、動画像データの内容確認を容易かつ適確に行える。また、本発明のデータ構造によれば、複数の画像処理装置間で、内容確認のための情報を共有することができる。

【図面の簡単な説明】

【図1】本実施形態の画像処理装置の制御構成を示すブロック図である。

【図2】本実施形態の画像処理装置の機能構成を示すブロック図である。

【図3】本実施形態の画像処理装置に格納されるシーン特徴データのデータ構造を示す図である。

【図4】本実施形態による画像のブロック分割例を示す図である。

【図5】本実施形態による多次元特徴量空間を説明する図である。

【図6】本実施形態による登録処理の手順を表すフローチャートである。

【図7】シーン特徴情報の一例を示す図である。

【図8】第1の実施形態によるダイジェスト動画作成処理の手順を表すフローチャートである。

【図9】第1の実施形態によるダイジェスト動画作成処理の手順を表すフローチャートである。

【図10】動画とその重要なシーンを概念的に表すとともに、これに対応させて各シーンの重要度と閾値の関係を概念的に表す図である。

【図11】本実施形態によるダイジェストシーン情報のデータ構成例を示す図である。

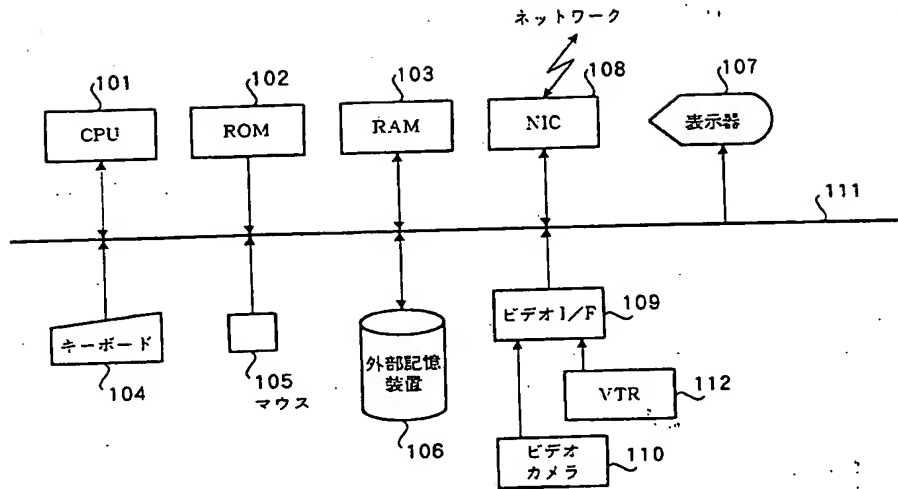
【図12】本実施形態による動画検索の際のユーザの典型的な操作における状態遷移図である。

【図13】本実施形態による動画検索の処理を説明するフローチャートである。

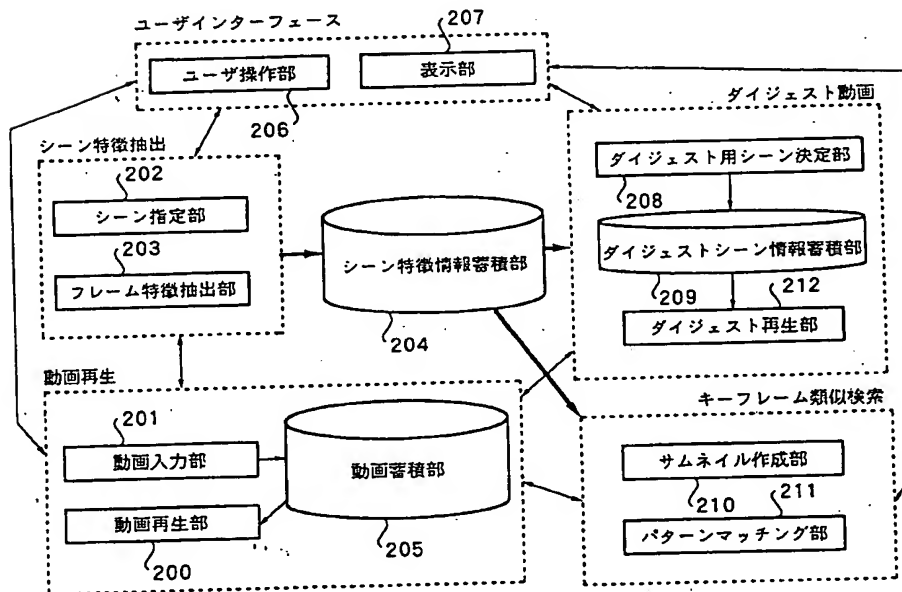
【図14】ラベル列間の類似度計算を行うためのペナルティマトリクスの例を示す図である。

【図15】第2の実施形態によるダイジェスト動画作成処理の手順を表すフローチャートである。

【図1】



【図2】



【図3】

(a) シーン特徴

開始フレーム番号
キーフレーム番号
終了フレーム番号
シーンの重要度
キーフレームのカラーレイアウト

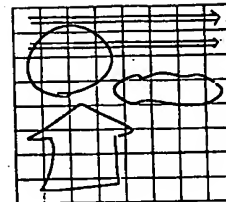
(b) キーフレームのカラーレイアウト

分割数 (2のべき乗)
色特徴の抽出方法
カラーデータフォーマット
ラベル列

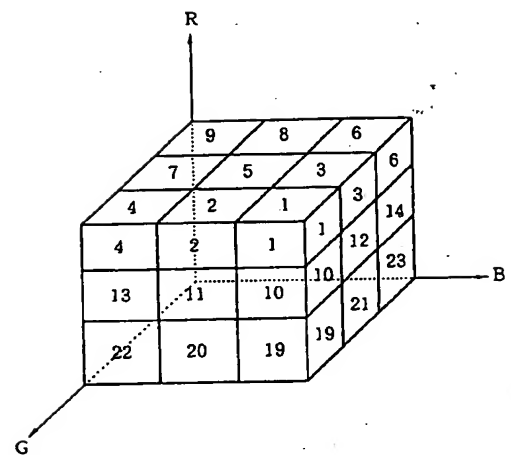
(c) カラーデータフォーマット

色空間
色の深さ
カラーマップ

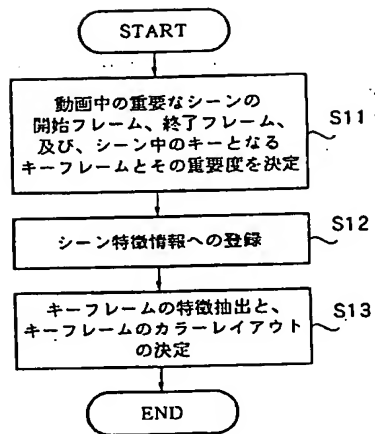
【図4】



【図5】



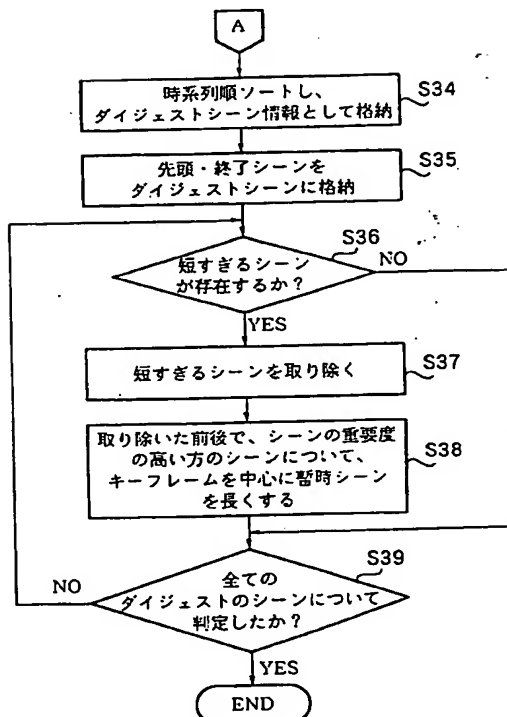
【図6】



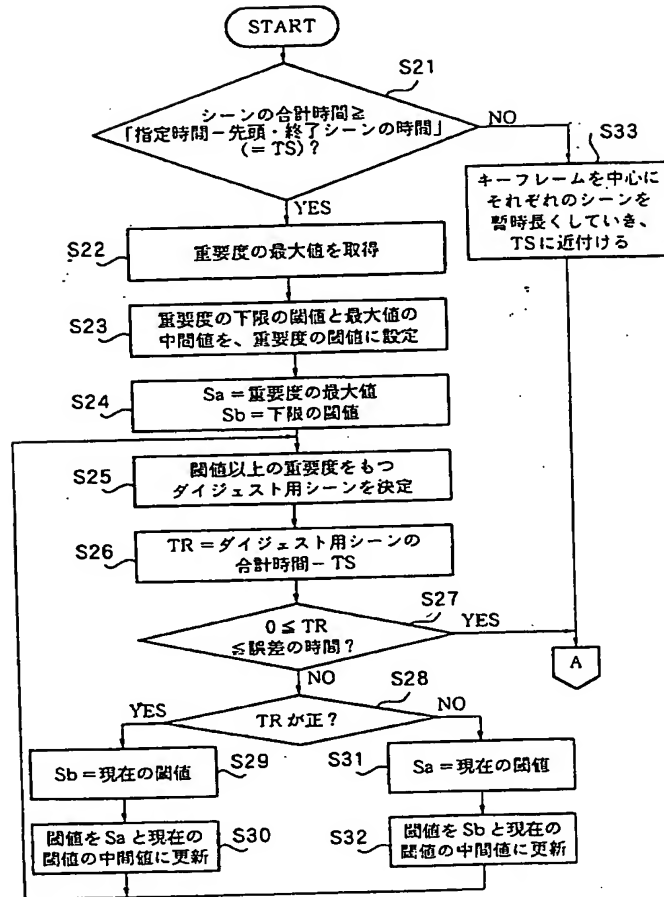
【図7】

動画ID	シーンID	シーン特徴
0	0	シーン特徴0
0	1	シーン特徴1
0	2	シーン特徴2
1	0	シーン特徴3
1	1	シーン特徴4
:	:	:

【図9】



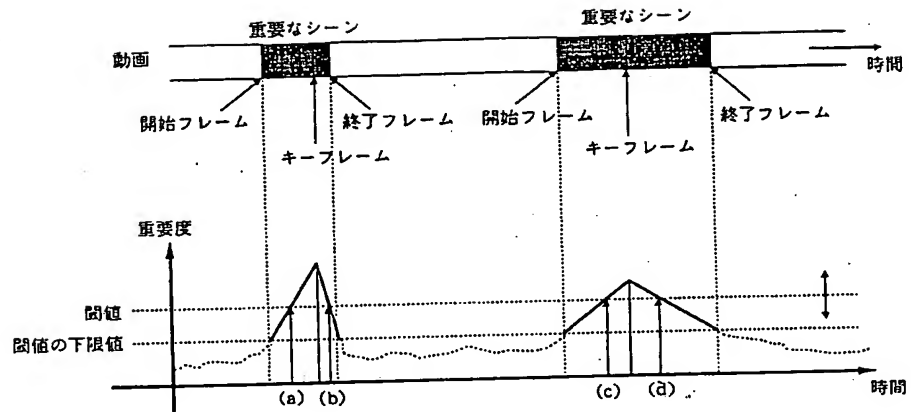
【図8】



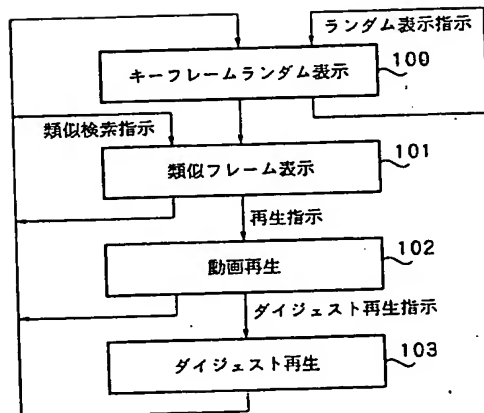
【図11】

動画ID	シーンID	開始フレーム番号	終了フレーム番号
0	0	50	55
0	1	110	125
0	2	300	308
0	3	411	416
1	0	20	25
⋮	⋮	⋮	⋮

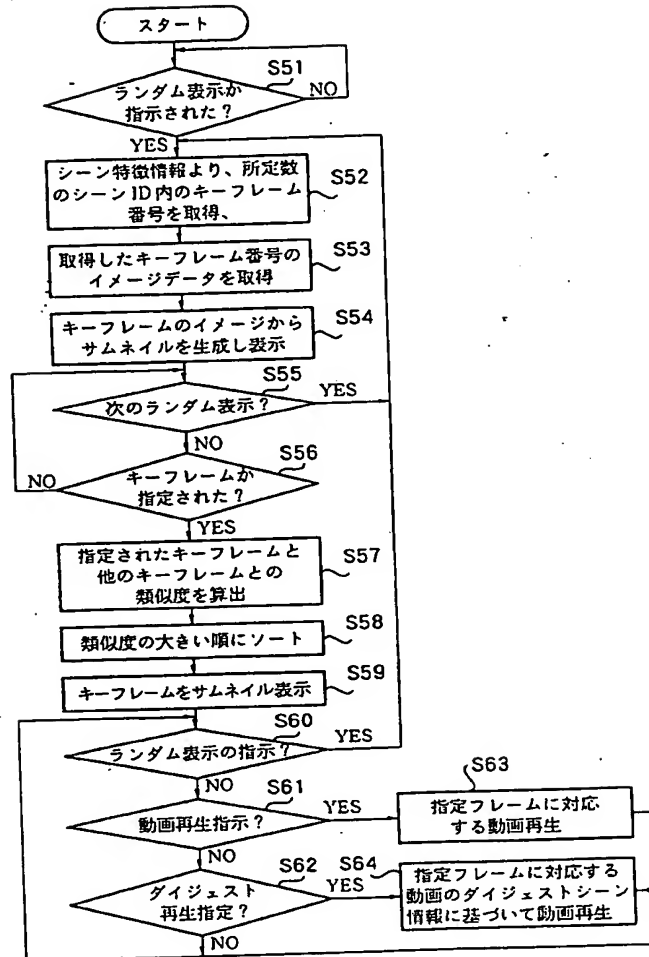
【図10】



【図12】



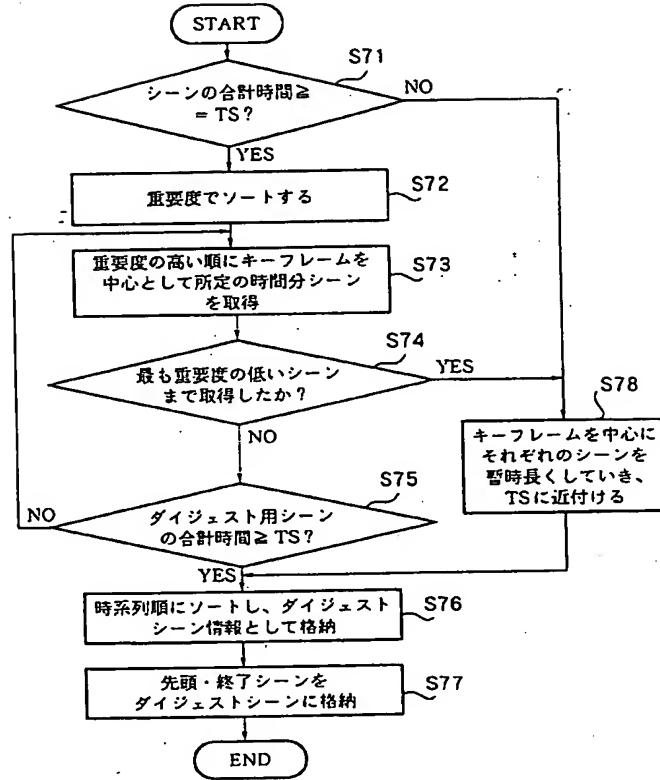
【図13】



【図14】

	1	2	3	4	5	6	7	8
1	0	1	1	5	1	5	7	7
2		0	2	1	1	7	2	5
3			0	7	1	1	5	2
4				0	1	9	1	7
:									

【図15】



フロントページの続き

(51) Int. Cl. 7

識別記号

F I
H O 4 N 5/91

テーマコード (参考)

Z

(72) 発明者 東條 洋

東京都大田区下丸子3丁目30番2号 キヤ
ノン株式会社内

F ターム (参考) 5B050 BA10 CA07 EA09 GA08
 5B075 ND12 NK06 PP02 PP03 PQ02
 PQ48 PR08 QM08
 5C052 AA01 AA16 AB04 AC08 CC06
 5C053 FA07 FA14 GB09 GB21 GB37
 HA29 HA30 JA21 JA24 LA06